# Venkata Shanmukha Sainath Gunda

+91 8688217161 venkatashanmukhasainathg@gmail.com LinkedIn GitHub Blog


## Education

**Indian Institute of Technology Kharagpur** **June 2019 – April 2023**

*Bachelor of Technology in Electronics and Electrical Communication Engineering; GPA: 8.8/10* *Kharagpur, West Bengal*

*Minor in Computer Science and Engineering, Micro in Artificial Intelligence and Applications*

## Experience

**KLA Corporation** **July 2023 – Present**

*Senior Analyst, Algorithm Research and Development* *Chennai, Tamilnadu*

- Developed ML algorithms for optical metrology, achieving angstrom-level precision in critical dimension measurement.
- Supported customers by developing features, optimizing algorithms, resolving bugs to enhance product performance.
- Developed an architecture to efficiently use synthetic data, improve robustness, and outperformed existing methods.

**KLA Corporation** **May 2022 – July 2022**

*Machine Learning Intern* *Chennai, Tamilnadu*

- Improved wafer defect classification F1-Macro by 3% on heavily imbalanced data using data-based approaches.
- Used Conditional GANs, Feature Generating Networks, SimCLR for data augmentation and representation learning.
- Applied Focal Loss, loss weighting, and sampling techniques to enhance model focus on minority defective classes.

**Amazon** **January 2022 – April 2022**

*Applied Scientist Intern*

- Enhanced image-based ad moderation models by integrating Vision Transformers, now deployed in production.
- Proposed a method to improve CLIP's few-shot learning using prompt tuning, cache model, and contrastive learning.
- Improved defect classification performance by 6% F1 Score in ad moderation through deep learning model optimization

## Projects

**AresSim** | Link **September 2026 - Present**

- Built a deterministic 32x32 Mars survival simumation from scratch to test and compare RL algorithms and LLM agents.
- Designed rover survival dynamics with ice, ore extraction, a 5x5 colony build, service, weather, and battery drain.
- Implemented a Gymnasium POMDP with 8x8 observations, action masking, masked PPO, and a play-replay UI.

**trisynapse-memory** | Link **August 2026 - Present**

- Built a local-first, evidence-first memory engine for AI agents with source-aware records and traceable citations.
- Developed hybrid multi-route retrieval combining BM25, vector embeddings, knowledge graph and temporal matching.
- Implemented complementary indexes: BM25, embeddings, knowledge graph, compiled claims, and group summaries.
- Fused route rankings with weighted RRF, then applied soft reranking, confidence gating, and context budgeting.

**Trisynapse** | Link **March 2026 - Present**

- Built a local-first research assistant converting scattered papers and feeds into structured, private knowledge vaults.
- Developed a multi-stage memory architecture with custom vector search and an interactive local semantic graph.
- Implemented background pipelines with local LLMs to autonomously monitor web sources and compile wikis.

**DotLM, SimpleThoughts & DotChat** | Link **February 2026 - April 2026**

- Trained **DotLM-165M**, a reasoning-capable language model from scratch in PyTorch using Qwen-like architecture.
- Created **Simple Thoughts**, a synthetic dataset created from LLMs spanning pretraining, SFT, DPO, and Reasoning.
- Tuned architecture and hyperparameters during training using Autoresearch, an agentic framework in a closed loop.
- Optimized GPU inference using CUDA Graphs, static KV caching, and FlashAttention to maximize token throughput.
- Developed **DotChat**, ChatUI featuring dynamic context gating via topic similarity to prevent context overflow.

**EssayInsightsAI** | Link **January 2025 - July 2025**

- Designed a discourse segmentation model by finetuning several BERT variants to segment essays into rhetorical elements
- Achieved 0.582 AUROC on Kaggle leaderboard using data-guided preprocessing, training, and post-processing methods
- Developed a Streamlit-based chatbot with OpenAI API, LangChain for real-time essay analysis, personalized feedback

**ScratchNLP** | Link **August 2024 - December 2024**

- Developed ScratchNLP, a Python library implementing foundational NLP algorithms from scratch using PyTorch.
- Implemented key algorithms, preprocessing methods, tokenization techniques, and metrics for various NLP tasks.
- Designed the project as a learning tool to enhance understanding of NLP concepts and their real-world applications.

**Bachelor's Thesis Project | Guide : Professor Jayanta Mukopadhyay** | Link **August 2022 - May 2023**

- Developed a Deep Learning and radiomics-based approach to predict MGMT promoter methylation in brain tumors.
- Combined segmentation and classification tasks using SwinUNETR and extracted Radiomics features for classification.
- Achieved a 0.609 AUROC on Kaggle leaderboard using various feature selection and ensemble learning methods

## Competitions / Hackathons

**Open Polymer Prediction 2025** | Link **July 2025 - September 2025**

- Designed property prediction or SMILES strings using ChemBERTa, GNNs, RDKit descriptors, and boosting models
- Achieved 0.097 weighted MAE using Optuna for hyperparameter optimization and SHAP for feature selection.

**Cloudphysician's Vital Extraction Challenge** **January 2023**

- Part of the silver medal-winning team in Cloudphysician's Vital Extraction Challenge of Inter IIT Tech Meet 2023
- Finetuned several OCR algorithms to detect vitals from patient monitor and achieved 77% accuracy with EasyOCR

**Cogito's IA for AI hackathon**| Link **September 2021**

- Developed an emotion classifier using OpenL3 embeddings and a KNN model trained on wav-converted speech data.
- Deployed the model as a Flask web app (HTML/CSS/JS frontend), yielding a top score of 58.876 in the hackathon.

**Amazon ML Challenge**| Link **August 2021**

- Classified multilingual product text (titles and descriptions) using translation and Sentence-Transformer embeddings
- Trained KNN ensembles with cross-validation achieving 66.85 % accuracy and ranked 22$^{nd}$ out of 3300 teams

**Techkriti's ML Hackathon**| Link **June 2021**

- Achieved 1$^{st}$ place by experimenting with several SOTA vision models to classify images of the Natural scenes dataset
- Trained DenseNet with regularization layers and augmentation, achieving 94.6% accuracy as the top-ranked solution.

## Technical Skills

**Languages / Skills**: Python, C++, Deep Learning, LLMs, Agentic AI, Reinforcement Learning, Computer Vision, NLP
**Technologies / Frameworks**: TensorFlow, PyTorch, PyTorch Lightning, Hugging Face, Gymnasium, LangChain, Scikit-Learn, OpenCV